

深層学習を用いた画像による作業員検出評価—追加学習による性能向上—

鹿島建設(株) 正会員 ○北原靖之 大塩 真 片村立太 露木健一郎 後閑淳司 中村泰広

1. はじめに

建設現場では時間外労働削減や生産性向上の観点から、施工管理の省力化技術が求められている。筆者らは監視カメラを活用した作業監視の省力化技術として深層学習型画像認識技術の適用を検討している。深層学習型モデルにおいては学習用のデータセットの数量を増やすことで画像認識の性能が向上することが広く知られているが、データセットの数量と性能の定量的な関係は目的のタスクや画像の条件に応じて大きく変化する。本稿では、現場で取得した画像から作成した作業員のデータセットを用いて追加学習を実施した際の性能向上について定量評価した結果を報告する。

2. 追加学習用データセットの作成とモデル訓練

本検討では、躯体構築における型枠組立、配筋組立、コンクリート打設の3工種・作業を対象とした。まずは作業現場において一日毎にアクションカメラ(GoPRO9)を設置し、作業の様子を捉えた動画を取得した。その後、作業動画からランダムにフレームを抽出し、専用ツールにて作業員の位置、サイズを明示することでデータセットを作成した。追加学習用画像データ例を写真-1に、データセットの条件を表-1に示す。検出モデルにはYOLOv5m²⁾を採用し、Microsoft COCOデータセットによる80クラス分類の学習済モデルに対して作成したデータセットによる追加学習を実施した。追加学習では分類クラス数を作業員1クラスに変更し、各ネットワーク層のパラメータは固定せずに訓練を実施した。



(a) 型枠組立 (A) (b) 鉄筋組立 (B) (c) コンクリート打設 (C)

写真-1 追加学習に使用した画像データの例(※掲載のため一部加工処理実施)

3. 検証内容と結果

表-1 作成した画像データセット条件

3.1 データセット数量と検出性能の関係

初めにデータセットの数量と検出性能向上の関係を定量評価するため、データセットAにおける訓練データ枚数を変化させたときの適合率、再現率の変化を評価した。結果を図-1に示す。適合率はモデルによる検出結果のうち、正しく作業員が検出された割合を示す指標である。また、再現率は画像に含まれる全ての作業員のうち、モデルにより正しく検出された割合を示す指標である。適合率、再現率共に200枚以降では単調増加の傾向となり1000枚で共に0.95程度の値となることが確認できる。また50枚は0枚(追加学習無し)よりも性能が低い結果となっており、訓練データセットが少量過ぎる場合は追加学習なしよりも性能が低下してしまう場合があることが確認できる。

	A	B	C
工種作業	型枠組立	鉄筋組立	コンクリート打設
訓練データ(枚)	1068	473	461
検証データ(枚)	150	150	150

キーワード 画像認識, 深層学習, 物体検出, 人体検出

連絡先 〒182-0036 東京都調布市飛田給 2-19-1 鹿島建設(株) 技術研究所 TEL 042-485-1111

3.2 データセットの組合せと性能指標の関係

次に訓練するデータセットの組合せを変えた場合の検出性能比較を実施した。訓練に用いる A,B,C の組合せは全部で8通り(追加学習無し, A, B, C, AB, BC, CD, ABC)あり, さらに検証に使用するデータセットが3通り(A, B, C)あるため, 訓練と検証のデータセットの全ての組合せには24通りのパターンが存在する。この全てのパターンに対してF値で評価を実施した。F値とは適合率, 再現率の平均を算出した指標である。結果を図-2に示す。

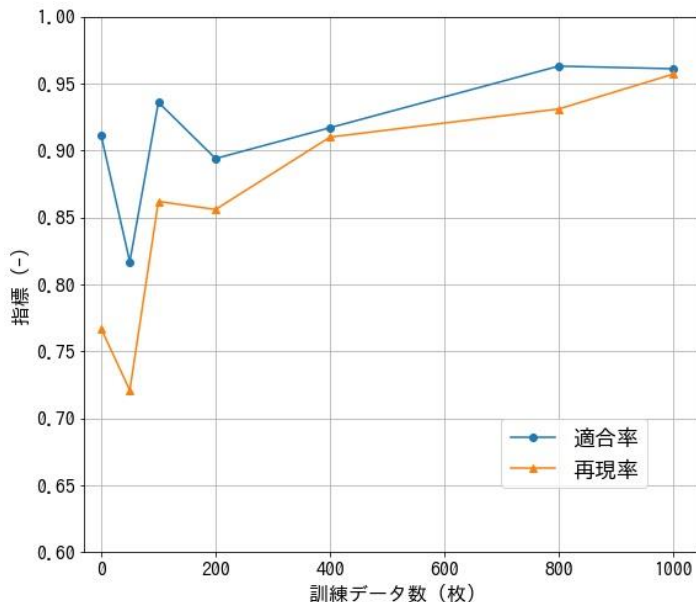


図-1 データセット A における訓練データ数量と性能指標との関係

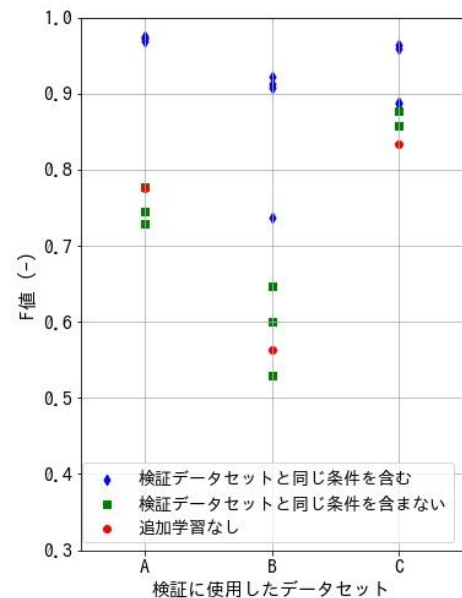


図-2 データセットの組合せと性能指標との関係

A, B, C いずれの条件においても, 訓練データに検証データセットと同じ条件を含む場合は, 高い F 値を示す場合が多いことが確認できる。これは訓練画像に含まれる特徴が効果的に学習されたことを示している。しかしながら, 訓練データに検証データセットと同じ条件を含まない場合は, 追加学習無しよりも高い値を取る点が存在する一方, 低い値を取る点も存在しており, 条件によっては追加学習により性能が低下してしまう場合があることが確認できる。この性能低下は訓練データの量が少なすぎたために発生したと考えられる。

工種毎の分布の傾向で比較すると, 性能の高い順にコンクリート打設(C), 型枠組立(A), 鉄筋組立(B)となった。コンクリート打設は見通しが良く, 人の認識が容易になるため, 検出性能が高くなりやすいと考えられる。また画像 1 枚当たりの作業員人数が多いことが再現率の向上に有利に働いた可能性がある。型枠組立は死角や遮蔽が多いため, 人の検出が難しいと考えられる。鉄筋組立は比較の見通しは良いものの, 今回作成したデータでは遠方に見かけのサイズの小さい作業員が映っており, それが要因で再現率が低下したと考えられる。

4. まとめ

深層学習型モデルによる作業員の検出性能に関して, 専用のデータセットを作成し, 躯体構築の3工種について定量評価した。その結果, 型枠組立の事例では1000枚程度の訓練により適合率, 再現率ともに0.95以上の値となることを確認した。また学習データセットの組み合わせを変えた評価を実施し, 評価データと同条件の訓練データを含めた方が性能が向上すること, 異なる条件の訓練データを用いる場合は追加学習により性能が低下する場合があることを確認した。さらに3工種の検出性能について比較, 考察を実施した。

今回の結果から作業員検出を目的とした深層学習モデルは追加学習による検出性能向上を図ることが可能であり, 作業監視の省力化技術として有用であることが確認できた。

参考文献

- 1) 北原ら: 深層学習モデルを用いた画像による作業員検出評価, VI-587, 土木学会全国大会第77回年次学術講演会, 2022.
- 2) YOLOv5: Glenn Jocher, <https://github.com/ultralytics/yolov5> (released 2020)